



Working Paper 06-15(04)
Statistics and Econometrics Series
February 2006

Departamento de Estadística
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (34) 91 624-98-49

PRINCIPAL ALARMS IN MULTIVARIATE STATISTICAL PROCESS CONTROL

Isabel González and Ismael Sánchez*

Abstract

This paper describes a methodology for the simulation of multivariate out of control situations using in-control data. The method is based on finding the independent factors of the variability of the process, and shifting these factors one by one. These shifts are then translated in terms of the observed variables. The shifts provoked by the most important factors are called principal alarms. The principal alarms are plotted, visualizing the main deviations of the process. Also, a resampling procedure for ARL estimation using principal alarms is proposed. An application using a real industrial process, illustrates the usefulness of the methodology.

Keywords: Average run length; Hotelling's T^2 ; MEWMA; Multivariate process control; Principal alarms;

* Gonzalez, Department of Mechanical Engineering, Universidad Carlos III de Madrid; Sánchez, Department of Statistics, Universidad Carlos III de Madrid, Avda. de la Universidad, 30, 28911 Leganés (Madrid), e-mail: ismael@est-econ.uc3m.es

Principal Alarms in Multivariate Statistical Process Control

Isabel González
Department of Mechanical Engineering
Universidad Carlos III de Madrid

Ismael Sánchez*
Department of Statistics
Universidad Carlos III de Madrid

February 27, 2006

Abstract

This paper describes a methodology for the simulation of multivariate out of control situations using in-control data. The method is based on finding the independent factors of the variability of the process, and shifting these factors one by one. These shifts are then translated in terms of the observed variables. The shifts provoked by the most important factors are called principal alarms. The principal alarms are plotted, visualizing the main deviations of the process. Also, a resampling procedure for ARL estimation using principal alarms is proposed. An application using a real industrial process, illustrates the usefulness of the methodology.

Key Words: Average run length; Hotelling's T^2 ; MEWMA; Multivariate process control; Principal alarms;

1 Introduction

The quality assessment of a production process is often attained by the joint monitoring of several correlated characteristics. It is of course possible to monitor the individual characteristics using univariate control charts. However, to the extent that the characteristics are correlated, a multivariate monitoring of the vector of characteristics can be more effective. A number of multivariate control charts have been proposed in statistical literature. One common method of constructing multivariate control charts is based on Hotelling's T^2 statistics (Hotelling, 1947; Alt, 1985). This T^2 chart can be considered as the multivariate extension of the univariate Shewart control charts, based on the monitoring of the means in independent samples. Multivariate CUSUM charts have been proposed by Woodall and Ncube (1985), and Croisier (1988). A multivariate extension of the exponentially weighted moving average (MEWMA), can be found in Lowry et al. (1992).

The use of multivariate control charts is, however, subject to practical drawbacks. The main problem of the multivariate control chart is the interpretation, in terms of the original quality variables, of an out of control signal. As a result, further diagnostic detective work would be needed in order

*Correspondence author, Universidad Carlos III de Madrid. Avd. de la Universidad 30, 28911, Leganés, Madrid (Spain), ismael.sanchez@uc3m.es, fax: 34-916249430

to know which of the original variables caused the alarm signal. This problem of interpreting an out of control situation in multivariate control charts is an important issue. Failure to quickly find the correct cause of the alarm makes the control impractical.

The identification of the cause of the alarm can be a complex issue, since it can be related to a single variable or to a subset of variables. This identification problem has been addressed by a number of authors. Murphy (1987) proposes a partitioning procedure using the overall T^2 of the observation that triggers the alarm, and comparing it to a T^2 value based on a subset of variables. The best subset is then searched based on this criterion. Doganaksoy et al. (1991) propose ranking the variables according to their relative contribution to the alarm using a univariate t statistic as a criterion. Hawkins (1993) and Wade and Woodall (1993) use regression adjustments for individual variables to help identify the single variable that is responsible of a the mean shift. Mason et al. (1995,1997) use some orthogonal decomposition of the T^2 value of the out of control observation to lead the interpretation.

It is important to note that the above mentioned procedures are based on the analysis of the T^2 statistics when the process is already out of control. Before the alarm, these procedures do not supply information about potential alarms. In this article, we propose a procedure aimed at identifying the main changes in the mean of the process that are expected to be observed in an out of control situation. These main changes in the mean of the process will be denoted as principal alarms. These principal alarms intend to summarize the most likely alarms that will appear in a multivariate control chart. Once the principal alarms are identify, it can be easier to find the physical causes of those alarms. Then, when a multivariate control chart triggers an alarm, we can start the identification of the causes using these principal alarms. The principal alarms are obtained by simulating alternative vector mean shifts so that they preserve the covariance structure of the observed data. The use of the proposed principal alarms analysis can be seen as an ex ante analysis of the out of control situation, whereas the traditional use of the out of control signals can be seen as an ex post analysis. Both approaches are complementary and help to make an effective statistical control.

The outline of the article is as follows. Section 2 describes the procedure to obtain the principal alarms of a multivariate process. Section 3 proposes a resampling procedure that uses the principal alarms to simulate the performance of alternative multivariate control charts. This procedure helps to select the best multivariate control chart. Section 4 shows an application of the proposed procedures to a real data set. Finally, Section 5 discusses some concluding remarks.

2 Simulation of Principal Alarms

2.1 General Considerations

In this section we describe the procedure to simulate the so called principal alarms. Let $\mathbb{X} = (X_1, X_2, \dots, X_K)'$ be the vector of the K variables of interest that are used for statistical process

control. Let \mathbf{X} be the $n \times K$ matrix of observed data. The vector of n observations of variable k is denoted as $\mathbf{x}_k = (x_{k1}, \dots, x_{kn})'$. Let us assume that $E(\mathbb{X}) \equiv \boldsymbol{\mu} = (\mu_1, \mu_2, \dots, \mu_K)'$ and $\text{Cov}(\mathbb{X}) = \boldsymbol{\Sigma}$.

One goal in statistical process control is to detect a multivariate mean shift $\boldsymbol{\mu}^* = \boldsymbol{\mu} + \boldsymbol{\delta}$, but the covariance $\boldsymbol{\Sigma}$ remains constant. The vector $\boldsymbol{\delta} = (\delta_1, \delta_2, \dots, \delta_K)'$ represents the shift in each original variable. The purpose of a multivariate control chart is to detect that such deviation $\boldsymbol{\delta}$ has taken place. Once a deviation is detected, the next goal is to investigate which elements of the vector $\boldsymbol{\delta}$ holds $\delta_k \neq 0$, and interpret it. Our goal here is different in that the interest is not in the analysis of observed mean shifts, but in the identification of potential mean shifts.

In order to investigate potential mean shifts, we should not think that each mean μ_k changes in an independent fashion. For instance, let \mathbb{X} represent a vector of measurements of a conformed steel profile at K selected points. The production process itself will make those measurements dependent. Then, a deviation in the conformation process will probably provoke a shift $\boldsymbol{\delta}$ in all the K measurements. As a result, the individual shifts δ_k will be dependent, since they are physically related. Then, in order to simulate a realistic mean shift we can not propose a deviation vector $\boldsymbol{\delta}$ with uncorrelated elements δ_k . We should keep in mind that practitioners will use multivariate control charts instead of a set of univariate ones just because the variables of interest are closely related. We should therefore design procedures for that context.

To simulate a realistic mean deviation, we can translate the problem of a mean shift in the set of related variables \mathbb{X} to a problem of a mean shift in a set of independent variables $\mathbb{Y} = (Y_1, Y_2, \dots, Y_J)'$, where J could be equal to K . This vector \mathbb{Y} can be interpreted as the original or primary independent sources of variability. Then, the vector \mathbb{X} is just a set of measurable magnitudes whose variability comes from \mathbb{Y} (apart from some specific variability like measurement errors).

The vector \mathbb{Y} could be obtained using an appropriate transformation $\mathbb{Y} = f(\mathbb{X})$. Since \mathbb{Y} is independent, we can easily simulate a mean shift just by adding a vector of constants $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_J)'$; where now the individuals γ_j , $j = 1, \dots, J$ do not need to be related between each other. This change will provoke a mean shift in \mathbb{Y} . Then, to translate this mean shift in terms of the original variables \mathbb{X} we just have to apply the inverse of the transformation, $f^{-1}(\mathbb{Y} + \boldsymbol{\gamma})$ and obtain $\boldsymbol{\delta}$ such that $\mathbb{X} + \boldsymbol{\delta} = f^{-1}(\mathbb{Y} + \boldsymbol{\gamma})$. Hence $\boldsymbol{\delta} = E[f^{-1}(\mathbb{Y} + \boldsymbol{\gamma})] - \boldsymbol{\mu}$.

In the literature there are alternative procedures to project a set of dependent variables \mathbb{X} into a set of independent ones \mathbb{Y} like, for instance, Independent Component Analysis (see, i.e. Hyvärinen et al, 2001). Under normality, a simpler procedure is the principal component analysis (PCA). Our proposed procedure will use PCA to find the transformation $\mathbb{Y} = f(\mathbb{X})$. We will not assume any particular distribution for \mathbb{X} . However, we will assume that \mathbb{X} is some transformation $\mathbb{X} = h(\mathbb{W})$, where \mathbb{W} are the original non-normal variables, so that \mathbb{X} is normally distributed. In practice, some Box-Cox type transformation can help to improve the normality of the data. In our context, since we are interested in the mean of the variables, a convenient transformation to improve the normality could be grouping the data in subsamples of size m . Then, instead of working with a set of n observations,

we would work with a set of $[n/m]$ observations. The advantage of this transformation is that the mean of the grouped variables is the same as the original ones. As a consequence, a mean shift in the grouped variables is the same as in the original variables. In the rest of the article, and in order to avoid further notation, we will assume that \mathbb{X} is normally distributed.

2.2 Computation of Principal Alarms

Let \mathbf{X} be a $n \times K$ matrix of observations of \mathbb{X} . Assuming that the process has been in control during these observations, they can be used to obtain the estimates $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$ of the mean and covariance matrix of \mathbb{X} respectively. We can also estimate the correlation matrix of \mathbb{X} , that will be denoted as $\hat{\mathbf{R}}$. To obtain the transformation $\mathbb{Y} = f(\mathbb{X})$ mentioned in the previous section, we could use the PCA based on the singular value decomposition of $\hat{\boldsymbol{\Sigma}}$ or $\hat{\mathbf{R}}$. We will develop here the procedure using $\hat{\mathbf{R}}$. The decomposition using $\hat{\boldsymbol{\Sigma}}$ follows the same arguments. The singular value decomposition of $\hat{\mathbf{R}}$ holds

$$\hat{\mathbf{R}} = \mathbf{C}\mathbf{D}\mathbf{C}', \quad (1)$$

where \mathbf{C} is the matrix with the eigenvectors of $\hat{\mathbf{R}}$ and \mathbf{D} is a diagonal matrix with the corresponding eigenvalues λ_k , $k = 1, \dots, K$. Hence, the transformed data \mathbf{Y} can be obtained as

$$\mathbf{Y} = \tilde{\mathbf{X}}\mathbf{C}, \quad (2)$$

where $\tilde{\mathbf{X}}$ is the standardization of \mathbf{X} , that is,

$$\tilde{\mathbf{X}} = (\mathbf{X} - \mathbf{1}\hat{\boldsymbol{\mu}}') \text{diag}(\hat{\boldsymbol{\Sigma}})^{-1/2}, \quad (3)$$

where $\mathbf{1}$ is a vector of ones of appropriate dimension, and $\text{diag}(M)$ is the matrix with the diagonal elements of M .

Under normality of \mathbb{X} , the transformed variables \mathbb{Y} are independent. Once we have the $(n \times K)$ values \mathbf{Y} we can introduce some mean shift in each column of \mathbf{Y} independently. The first column of \mathbf{Y} , denoted as \mathbf{y}_1 , is the linear combination with maximum variance. The corresponding eigenvalue λ_1 is its variance. As it is well known, and using that $\sum \lambda_k = K$, the larger the correlation between the original variables \mathbf{X} the larger λ_1 , and hence, the smaller λ_k , $k > 1$. As a result, if we are dealing with a set of correlated variables, the first principal components can be seen as the main sources of variability of the process. For instance, if $K = 2$, we have from (2) that

$$\mathbf{y}_1 = c_{11}\tilde{\mathbf{x}}_1 + c_{21}\tilde{\mathbf{x}}_2,$$

where c_{ij} is the (i, j) element of the matrix \mathbf{C} , and $\tilde{\mathbf{x}}_k$ is the k -th column of $\tilde{\mathbf{X}}$. From this expression, we can write the variance of \mathbf{y}_1 as

$$\lambda_1 = \frac{\mathbf{y}_1' \mathbf{y}_1}{n} = c_{11}^2 \frac{\tilde{\mathbf{x}}_1' \tilde{\mathbf{x}}_1}{n} + c_{21}^2 \frac{\tilde{\mathbf{x}}_2' \tilde{\mathbf{x}}_2}{n} + 2c_{11}c_{21} \frac{\tilde{\mathbf{x}}_1' \tilde{\mathbf{x}}_2}{n}. \quad (4)$$

Since the eigenvectors have unity norm and $\tilde{\mathbf{x}}_k$ are standardized, we obtain from (4) that

$$\lambda_1 = 1 + 2c_{11}c_{21}r, \quad (5)$$

where r is the correlation between \mathbf{x}_1 and \mathbf{x}_2 . Since $\lambda_1 \geq \lambda_2$ and $\lambda_1 + \lambda_2 = 2$, it holds from (5) that $c_{11}c_{21}r \geq 0$. Then, it can be checked that as $|r| \rightarrow 1$ then $\lambda_1 \rightarrow 2$ and $\lambda_2 \rightarrow 0$. The interpretation of this result is that the first columns of \mathbf{Y} are a good approximation of the primary independent sources of variability of \mathbf{X} . Therefore, if we want to predict a realistic mean shift in the process, it suffices to simulate a mean shift in the first columns of \mathbf{Y} . As a consequence, the resulting multivariate mean shift in terms of the \mathbb{X} , that is a shift $\boldsymbol{\mu}^* = \boldsymbol{\mu} + \boldsymbol{\delta}$, will be coherent with the structure of the data and will have practical meaning.

Let us denote as \mathbf{Y}^* to the matrix obtained after adding a constant term in some columns of \mathbf{Y} ; that is

$$\mathbf{Y}^* = \mathbf{Y} + \mathbf{1}\boldsymbol{\gamma}'. \quad (6)$$

The shifted original variables are

$$\mathbf{X}^* = \mathbf{Y}^* \mathbf{C}' \text{diag} \left(\hat{\boldsymbol{\Sigma}} \right)^{1/2} + \mathbf{1}\hat{\boldsymbol{\mu}}'. \quad (7)$$

Since $\hat{\boldsymbol{\mu}} = n^{-1} \mathbf{1}' \mathbf{X}$, we can write the mean shift of the original variables as

$$\boldsymbol{\delta} = \frac{1}{n} (\mathbf{X}^* - \mathbf{X})' \mathbf{1}. \quad (8)$$

This shift in the process obtained from a shift in the principal components of \mathbf{X} will be denoted as principal alarms (PA). Note that a PA is a deviation in the process that has not been observed yet. Therefore, it is a piece of information that would be difficult to achieve just exploring the observed data. It will hence be very illustrative to make a graphical representation of the PA that compares $\boldsymbol{\mu}$ and $\boldsymbol{\mu}^*$. This representation will help us to understand the process better. Since the principal components are assumed to be independent, it will be more practical to simulate a PA shifting the principal components one by one. For instance, a PA of the first principal component will use $\boldsymbol{\gamma} = (\gamma_1, 0, \dots, 0)'$. This PA will be denoted as the first PA. Usually, in SPC, the mean shifts are defined in terms of the standard deviation of the process. In the present context, we can define the k -th PA as the one obtained with the shift

$$\boldsymbol{\gamma} = \left(0, \dots, 0, b\sqrt{\lambda_k}, 0, \dots, 0 \right)', \quad (9)$$

where b determines the size of the shift.

This procedure to simulate mean shifts using the PA can also be used to simulate shifts in the covariance matrix of the \mathbb{X} , maintaining the mean constant. In this case, instead of adding the constant $\gamma_k = b\sqrt{\lambda_k}$ to the k -th column of \mathbf{Y} , we will multiply that column by $\gamma_k = \sqrt{(1+b)}$. Then, the variance of that column will be $\text{var}(\mathbf{y}_k^*) = (1+b)\lambda_k$, whereas the mean remains constant. This shift in the variance of \mathbf{y}_k will change the whole covariance matrix of \mathbf{X} according to the distortion

introduced by γ_k . Note that, as also exposed in the mean shift case, a realistic change in the covariance matrix of \mathbf{X} can not be attained by changing the elements of $\hat{\Sigma}$ arbitrarily. If we still assume that the principal components \mathbb{Y} are the independent and original sources of variability of our production process, then changes in the variance of such original sources should also be considered independently. Therefore, realistic changes in Σ (or \mathbf{R}) that have practical meaning should be restricted to those that maintain its eigenvectors. A change in the eigenvectors is only possible if a radical and drastic change has happened in the process. For instance, the independent sources of variability become dependent. It does not seem reasonable to design control charts based on those situations.

3 Estimation of the Average Run Length using PA

The PA can also be used to evaluate the performance of alternative multivariate control charts, helping to chose the most appropriate for our process. In order to evaluate such performance we will use the so called average run length (ARL). The ARL is the average number of samples needed to obtain an out of control signal. For a given in-control ARL (related to Type I error) the best control chart is the one that minimizes the out-of-control ARL.

This section shows a procedure to simulate the ARL curve of a multivariate control chart as a function of b in (9). The proposed methodology is based on a resampling scheme from the data that permit to easily reproduce both in control and out of control situations, when the out of control situation is due to a PA. The proposed procedure can be applied to any multivariate control chart. In this article, for convenience, we will expose the procedure for the T^2 and MEWMA charts.

Let us denote as $\tau(\theta)$ to the statistic represented in the multivariate control chart, where θ is the set of parameters of the multivariate distribution (μ or Σ) that we want to control. Let UCL and LCL be the upper control limit and lower control limit, respectively, of the control chart. Let us suppose that we are interested in monitoring μ . When the process is in control, the multivariate distribution of the K variables has $\mu = \mu_0$. When the process is out of control due to a mean shift, the multivariate distribution has $\mu^* = \mu_0 + \delta$. The mean shift δ due to the k -th PA is originated by a mean shift $\gamma = (0, \dots, 0, \gamma_k, 0, \dots, 0)$, as in (9) where k will usually be $k = 1$ or $k = 2$. Therefore we can write $\mu^* \equiv \mu^*(\gamma_k)$. Given a value of in-control ARL, the best control chart to detect such PA is the one with minimum $ARL(\mu^*(\gamma_k))$. For economy of notation, we will denote the ARL curve as $ARL(\gamma_k)$.

3.1 ARL for the T^2 chart

The T^2 chart displays the evolution of the Hotelling's T^2 statistics built using subgroups of g observations. We will denote as $\bar{\mathbf{x}}_t$ to the vector of sampling averages of the g observations at time t . Then, the T^2 statistics is

$$T_t^2 = g (\bar{\mathbf{x}}_t - \hat{\mu})' \hat{\Sigma}^{-1} (\bar{\mathbf{x}}_t - \hat{\mu}). \quad (10)$$

If we could assume that the vector $\bar{\mathbf{x}}_t$ is normal, the (asymptotic) control limits of the T^2 chart are $LCL=0$, and $UCL=\chi_{K,\alpha}^2$. Then, the ARL curve could be calculated analytically under the shifts $\boldsymbol{\mu}^*(\gamma_k)$ using the χ_K^2 distribution. Alternatively, the ARL curve can be obtained in a easy way by using resampling methods similar to those in the literature (Bagjier, 1992; Seppala et al. 1995; Liu and Tang, 1996; Jones and Woodall, 1998). The asymptotic correctness of the bootstrap applied to the empirical distribution of the T^2 statistics has been verified in Beran (1986).

The proposed multivariate resampling procedure to estimate the in-control ARL, that is $ARL(0)$, of the T^2 is the following:

1. Obtain the estimates $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$, using all the data (we should make sure that the process is in control).
2. Select randomly g observations from \mathbf{X} (with replacement), and compute $\bar{\mathbf{x}}_t$ and the T_t^2 statistics (10).
3. If $T_t^2 \leq h_1$, where $h_1 > 0$ is a specified UCL, we repeat step 2 extracting a new set of g observations. If $T_t^2 > h_1$ we stop and then compute the run length as $RL=t$.

The $ARL(0)$ is estimated by averaging the RL 's obtained after repeating the previous procedure a large number of B times. The value of h_1 is chosen to achieve a specified $ARL(0)$. Since in this T^2 chart, the T_t^2 are independent, it holds that Type I error (false alarm probability) is $\alpha = 1/ARL(0)$.

The procedure to estimate the out-of control ARL due to the k -th PA, that is $ARL(\gamma_k)$, is the following:

4. Select a desired value of b in (9) and compute γ . Then compute \mathbf{Y}^* as in (6), and \mathbf{X}^* as in (7).
5. Do steps 2 and 3 using \mathbf{X}^* instead of \mathbf{X} and obtain the RL .

In a similar fashion as in $ARL(0)$, the $ARL(\gamma_k)$ is estimated by averaging the RL 's obtained after repeating the previous procedure a large number of B times. An ARL curve can be draft using a set of alternative values of b . For instance, we could move b in the range $(0,3)$ to analyze a positive shift.

As mentioned in Section 2.1, in the case of non-normal data, the matrix \mathbf{X} would be the transformation $\mathbf{X} = h(\mathbf{W})$, with \mathbf{W} being the original data, and $h(\mathbf{W})$ some transformation aimed at reaching normality. In this case, the random selection in steps 2-3 above are made from \mathbf{W} instead of \mathbf{X} . The step 4 should add the computation $\mathbf{W}^* = h^{-1}(\mathbf{X}^*)$. Then, step 5 is made with \mathbf{W}^* instead of \mathbf{X}^* .

Since we are estimating an empirical percentile in the tail of a distribution, we should have a large enough dataset, especially if we are going to build a control chart for individual observations ($g = 1$). For instance, if we want to estimate the the usual percentile $\alpha = 0.0027$ ($ARL(0) = 370$) for an individual observations chart, we should have a dataset larger than $1000/2.7 = 370$ observations.

An alternative way to see this problem is in terms of the ARL. If we want to estimate the event so that the ARL is 370 by averaging empirical run lengths, we should have a dataset larger than 370.

Otherwise, we would be asking the data set about events that has not been observed yet. When we use a subgroup size $g > 1$, this problem is alleviated, and we can obtain results from resampling using smaller datasets. First, the averages of $g > 1$ observations have smaller variance. Hence, extreme percentiles will be closer to the mean (in terms of standard deviations). Also, with $g > 1$ we are increasing the number of different averages $\bar{\mathbf{x}}_t$'s obtained from resampling by the power of g , which also ease the estimation of extreme percentiles.

3.2 ARL for the MEWMA chart

The MEWMA chart is implemented as in Lowry et al. (1992). Lowry et al. (1992) compute the statistics

$$\mathbf{Z}_t = r(\bar{\mathbf{x}}_t - \hat{\boldsymbol{\mu}}) + (1 - r)\mathbf{Z}_{t-1}, \quad (11)$$

where $\mathbf{x}_t = (x_{1t}, x_{2t}, \dots, x_{kt})'$ is the vector of observations at period t , $\mathbf{Z}_0 = \mathbf{0}$, and $0 < r \leq 1$ is the parameter of the chart used to downweight the past. The MEWMA chart gives an alarm signal if

$$D_t^2 = \mathbf{Z}_t' \boldsymbol{\Sigma}_{\mathbf{Z}_t}^{-1} \mathbf{Z}_t > h_2, \quad (12)$$

where h_2 is and UCL chosen to achieve a specified $\text{ARL}(0)$; and $\boldsymbol{\Sigma}_{\mathbf{Z}_t}$ is the covariance matrix of \mathbf{Z}_t , that is,

$$\boldsymbol{\Sigma}_{\mathbf{Z}_t} = \left\{ r \left[1 - (1 - r)^{2t} \right] / (2 - r) \right\} \hat{\boldsymbol{\Sigma}} / g. \quad (13)$$

Since the sequence of D_t^2 is not independent, the ARL should be obtained using a variation of the procedure used for the T^2 chart. The proposed procedure is the following:

- 1'. As in step 1, we obtain the estimates $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$, using all the data (assuming that the process is in control).
- 2'. In order to assure a steady-state behavior of D_t^2 , we extract an initial set of N subgroups of size g , where for each subgroup we extract randomly g observations from \mathbf{X} (with replacement). Then, for each subgroup, compute $\bar{\mathbf{x}}_s$, $s=1, \dots, N$, and \mathbf{Z}_s , using $\mathbf{Z}_0 = \mathbf{0}$. We set $t = 1$.
- 3' Select randomly g observations from \mathbf{X} (with replacement), and compute $\bar{\mathbf{x}}_{N+t}$, and \mathbf{Z}_{N+t} , and the D_t^2 statistics as

$$D_t^2 = \mathbf{Z}_{N+t}' \boldsymbol{\Sigma}_{\mathbf{Z}_{N+t}}^{-1} \mathbf{Z}_{N+t} \quad (14)$$

- 4'. If $D_t^2 \leq h_2$, we repeat step 3' extracting a new set of g observations ($t = t + 1$). If $D_t^2 > h_2$ we stop and then compute the run length as $\text{RL}=t$.

The $\text{ARL}(0)$ is estimated by averaging the RL's obtained after repeating the previous procedure a large number of B times. Note that, since the D_t^2 are not independent, then $\alpha \neq 1/\text{ARL}(0)$.

The procedure to estimate the out-of control ARL of the MEWMA chart due to the k -th PA, $\text{ARL}(\gamma_k)$, is the following:

5'. Repeat steps 1' and 2' and obtain $\bar{\mathbf{x}}_N$, and \mathbf{Z}_N .

6'. As in step 4, select a value of b in (9) and compute γ . Then compute \mathbf{Y}^* as in (6), and \mathbf{X}^* as in (7).

7'. Do steps 3' and 4' using \mathbf{X}^* instead of \mathbf{X} and obtain the RL.

The $\text{ARL}(\gamma_k)$ is estimated by averaging the RL's obtained after repeating the previous procedure a large number of B times. In EWMA type charts, it is easier to build charts based on resampling from a dataset of moderate or small size than in T^2 type charts. In order to see this point, we can rewrite expression (11) as

$$\mathbf{Z}_t = r \sum_{s=0}^t (1-r)^s (\bar{\mathbf{x}}_{t-s} - \hat{\boldsymbol{\mu}}), \quad (15)$$

which shows that the information of s periods ago is downweight by the factor $(1-r)^s$. From (15), it can be interpreted that the effective number of observations used in the computation of \mathbf{Z}_t is, asymptotically,

$$N_t = \frac{1}{r}.$$

Then, with the usual values of r (like $r = 0.1$) EWMA charts can be seen as an average of N_t observations. As a result, using (13), it can be verified that the smaller the parameter r the smaller the variance of \mathbf{Z}_t . Extreme percentiles of \mathbf{Z}_t will then be closer to its mean than the extreme percentiles of \mathbf{X} from its mean. Therefore, if the dataset is not large enough, we can still estimate an empirical percentile for $\text{ARL}(0) = 370$ if r is small enough.

4 Application to real data

4.1 General description

In this section we will apply the proposed methodology to analyze the PA using data from a real production process. The process is the manufacturing of the window frame for the door of a vehicle. The shape of this frame has very restrictive dimensional tolerances. On the one hand, this window frame must be fitted into the bodywork. On the other hand, it should allow the allocation of the window. Figure 1 shows this frame. Figure 1 (a) shows the profile of the frame where the 7 circles represent the 7 points that are measured to control the dimensions. The magnitudes of interest (among others) are the gaps between the frame and a special gauge built for this purpose. Figure 1 (b) shows the window frame fitted in this gauge. The gauge has a set of 7 holes that allow the introduction of a measurement instrument to measure the 7 distances to the frame.

We have a total of 250 measured frames. After some statistical analysis, 16 frames were discarded for being outliers, which leave with a total of 234 available frames. We will denote as \mathbb{W} to the set of 7 variables representing the 7 measured points and \mathbf{W} to the matrix of 234×7 measurements. The

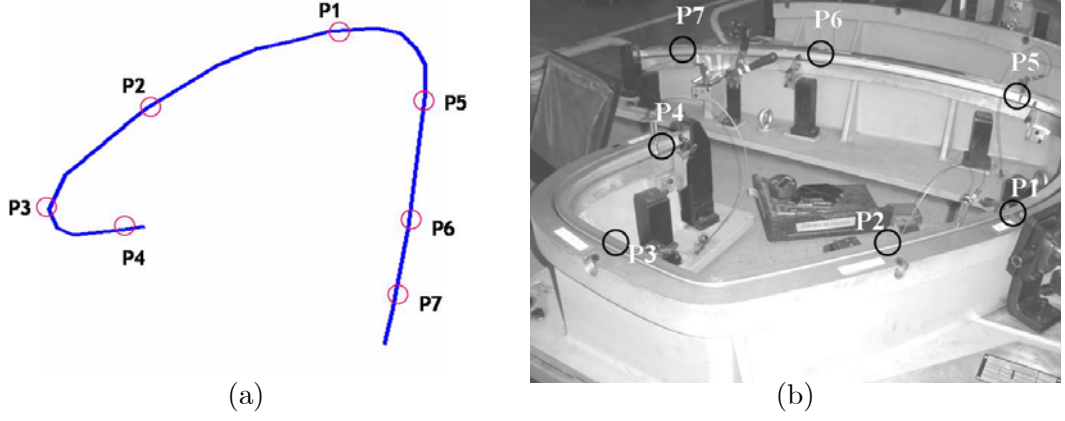


Figure 1: Window frame of the front door of a vehicle. (a) Profile with the seven measured control points. (b) Picture of the window frame fitted in the gauge used to measure the seven control points.

analysis of \mathbf{W} revealed that the variables were not normally distributed, due to different patterns of asymmetry. In order to improve the normality of the data, we have grouped the data in subgroups of size $m = 3$ and computed the mean of each subgroup. We denote as \mathbb{X} to the set of 7 transformed variables and \mathbf{X} to the matrix of 78×7 observed means. An analysis of univariate normality of each column of \mathbf{X} by Kolmogorov-Smirnov did not reject normality. Also, an analysis of multidimensional normality of \mathbf{X} , based on the multivariate skewness and the multivariate kurtosis, did not reject the normality assumption (at 5% significance level). Note that $E(\mathbb{X}) = E(\mathbb{W})$ and $\text{cov}(\mathbb{X}) = \text{cov}(\mathbb{W})/m$. Therefore, it is easy to analyze changes in the mean and covariance of \mathbb{W} using \mathbb{X} .

Since the 7 variables are of similar nature and similar order of magnitude, the PCA of \mathbf{X} has been made using the sampling covariance matrix of \mathbf{X} , denoted as $\hat{\Sigma}$. The matrix \mathbf{C} of eigenvectors of $\hat{\Sigma}$ holds $\hat{\Sigma} = \mathbf{C}\mathbf{D}\mathbf{C}'$. The first three principal components account for 70% of the variability in \mathbf{X} . The first component accounts for 39% of the variability, the second one 19% and the third one 12%. In order to obtain the principal alarms, we will modify only these three principal components.

4.2 Estimating mean shifts with principal alarms

In this section we will generate the first three PA by modifying each one of the first three principal components, respectively. Each principal component is modify by adding a shift γ_k , $k = 1, 2, 3$, respectively, as in expression (9), with b ranging in the interval $[-2, 2]$. Then, the shift is expressed in number of standard deviations of each component. Let \mathbf{Y} be the 78×7 matrix of projections on the principal components. That is, $\mathbf{Y} = \tilde{\mathbf{X}}\mathbf{C}$; where now $\tilde{\mathbf{X}} = (\mathbf{X} - \mathbf{1}\hat{\mu}')$ and \mathbf{C} is the matrix of eigenvectors of $\hat{\Sigma}$. To generate the first PA, we change the first column of \mathbf{Y} adding a shift $\gamma_1 = b\sqrt{\lambda_1}$, where $\lambda_1 = 0.176$, and the remaining columns of \mathbf{Y} are unaltered. This transformation produces \mathbf{Y}^* , which leads to $\mathbf{X}^* = \mathbf{Y}^*\mathbf{C}' + \mathbf{1}\hat{\mu}'$. Then, the sampling average of the columns of \mathbf{X}^* are the new means

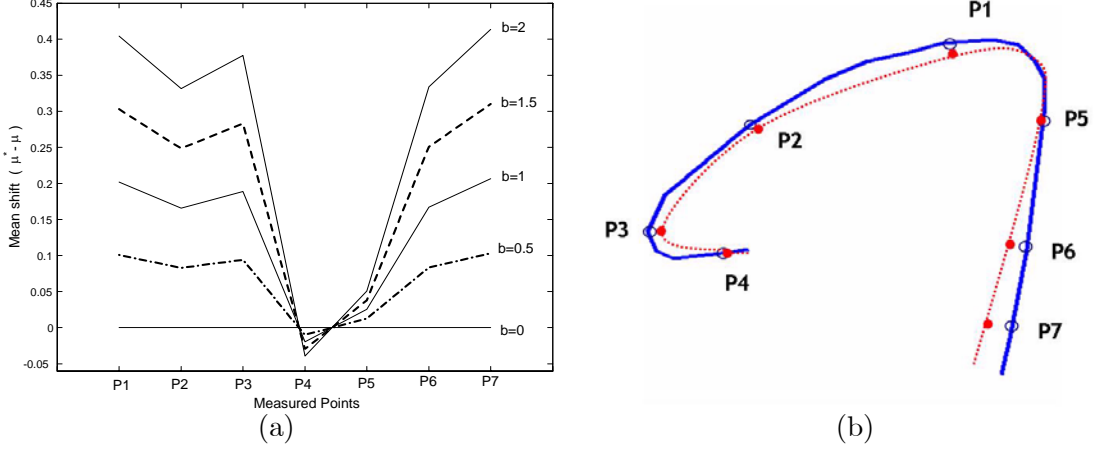


Figure 2: Simulated mean shifts provoked by the first PA in each measured point for different values of b . (a) Mean shift at each point. (b) is a simulated deformations (dotted line) from the nominal frame (solid line) provoked by a large value of $b > 0$.

$\hat{\mu}^*$.

We will make a graphical representation of the frame with the distorted $\hat{\mu}^*$. This graphical representation will be a useful tool to the quality practitioner to understand the main potential deformations of the window frame, and therefore to design a more efficient management system of the manufacturing process.

Figure 2 (a) shows the mean shift $\hat{\mu}^* - \hat{\mu}$ obtained as a function of b , for some values of $b \geq 0$. For $b < 0$ the mean shifts are the same but with opposite sign. We see in this figure that the main shifts due to the first PA will be seen in the points P1, P3, and P7, whereas P4 and P5 will have very little distortion. Figures 2 (b) displays the original mean frame (solid line) with the simulated deformed frame using a large value of $b > 0$. This figure suggests that the first PA will be produced by a wrong curvature in the mechanical bending between the points P1 and P5.

To generate the second PA, we change the second column of \mathbf{Y} adding a shift $\gamma_2 = b\sqrt{\lambda_2}$, where $\lambda_2 = 0.086$, and proceed as in the first PA. Figure 3 (a) shows the mean shift $\hat{\mu}^* - \hat{\mu}$ obtained as a function of $b > 0$. We see in this figure that the main shifts due to the second PA will be seen in the points P1 and P6. Figure 3 suggests that the second PA will be produced by a wrong curvature in the right side of the profile, between points P5 to P7, that will alter the position of the remaining points accordingly, especially P1.

To generate the third PA, we change the third column of \mathbf{Y} adding a shift $\gamma_3 = b\sqrt{\lambda_3}$, where $\lambda_3 = 0.053$, and proceed as in the previous PA. Figure 4 (a) shows the mean shift $\hat{\mu}^* - \hat{\mu}$ obtained as a function of b . We see in this figure that the main shifts due to the third PA is in P3. Figure 4 suggests that the third PA will be produced by a wrong mechanical bending of the elbow formed at P3.

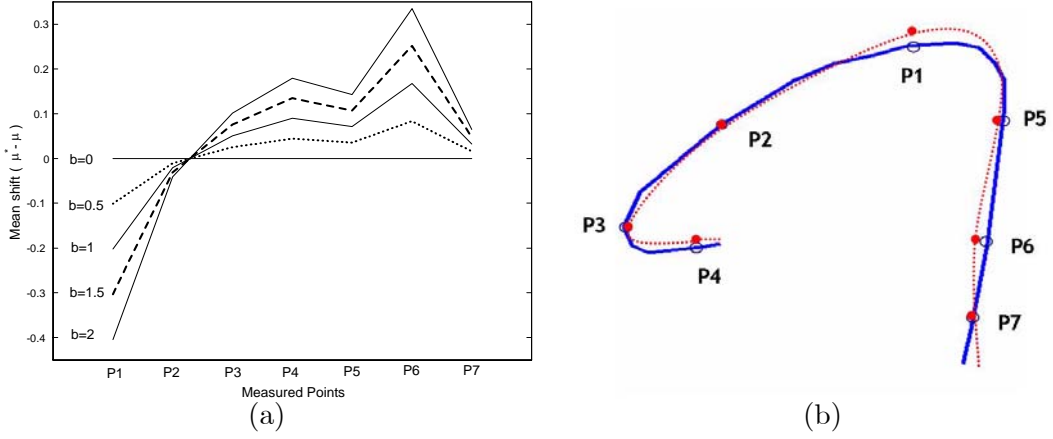


Figure 3: Simulated mean shifts provoked by the second PA in each measured point for different values of b . (a) Mean shift at each point. (b) is a simulated deformations (dotted line) from the nominal frame (solid line) provoked by a large value of $b > 0$.

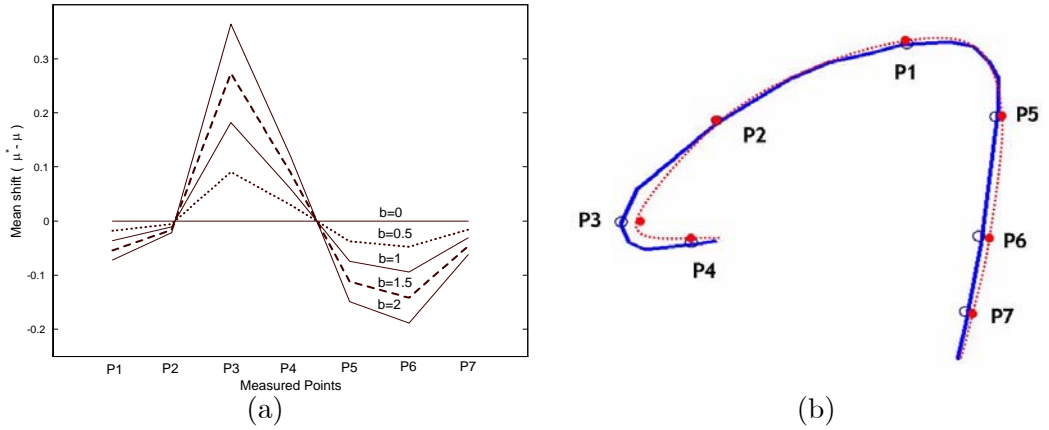


Figure 4: Simulated mean shifts provoked by the third PA in each measured point for different values of b . (a) Mean shift at each point. (b) is a simulated deformations (dotted line) from the nominal frame (solid line) provoked by a large value of $b > 0$.

4.3 Selection of the best multivariate control chart

In this section we will obtain the ARL curves for T^2 and MEWMA. The curves describe the ARL when a mean shift provoked by the first PA takes place. The ARL curves are built using the methodology proposed in Section 3. The comparison of the curves of the two charts will help to select the best multivariate control chart for real time operation of this process. This methodology can also be applied to the shifts provoked by other PA's. For the sake of brevity, the analysis of the second and third PA's are omitted.

We will analyze the performance of three alternative charts. The first is a T^2 chart implemented as in (10) using subgroup size $g = 3$. The other two charts are MEWMA charts implemented using $r = 0.10$, $N=1000$ initial subgroups, and subgroup sizes $g = 1$ and $g = 3$, respectively. In each of

these three competing charts, the control limits h_1 and h_2 have been obtained to achieve an in-control $ARL(0) = 370$. This ARL is the usual value for T^2 chart, which corresponds to a Type I error of $\alpha = 0.027$ in this chart. The values of h_1 and h_2 are obtained after some search. This search has been initiated using the value $\chi_{7;0.9973}^2$ as initial value for h_1 and h_2 . Then, these values are modified according to the resulting ARL. As in Croiser (1988) and Lowry et al. (1992), the search of the appropriate h_1, h_2 that produces $ARL(0) = 370$ can be helped using the regressing of $\log(ARL)$ on $h_i, i = 1, 2$. The ARL's for different h_i 's are obtained with the proposed resampling procedure using $B=100,000$ replications. Once h_1 and h_2 are established, we estimate each $ARL(\gamma_1)$ using $B=10,000$ replications.

Figure 5 (a) shows the ARL curves for the first PA as a function of b , where the shift of the first PA is $\gamma_1 = b\sqrt{\lambda_1}$. This figure reveals that the MEWMA charts are more efficient than the T^2 chart, and therefore they are the recommended ones. The decision of the value of g in the MEWMA chart depends on the inspection procedures. Figure 5 (a) shows that, as expected, the larger the subgroup size g the better the ARL. In the present example, the measurement of the control points is costly and time consuming. The current inspection procedure is limited to one inspected article at each period. Then, if we want to build a chart using $g = 3$ we should wait three times as much, delaying the computation of the new point of the chart. Figure 5 (b) helps us to select g in this situation. This figure displays the values of $ARL \times g$, which represent the average number of inspected articles, also known as average time to signal. Figure 5 (b) shows that the recommended chart is then the MEWMA with $g = 1$. We can continue using this methodology to check alternative values of r and find the optimal one.

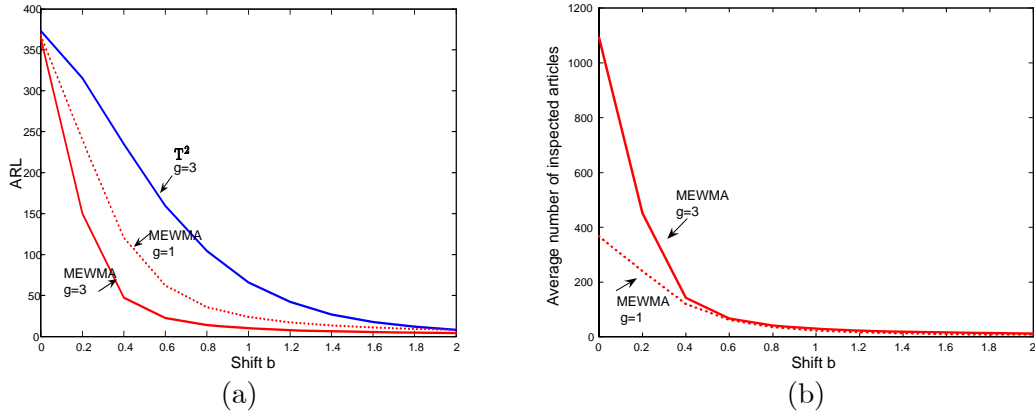


Figure 5: ARL curves of T^2 and MEWMA for the first PA. (a) Estimated ARL curves for different values of b and with $ARL(b=0)=370$. (b) Estimated number of inspected articles to get an out of control signal ($ARL \times g$).

5 Discussion

Traditionally, the identification of causes in multivariate control problems is made when the process is actually out of control. This practice can be harmful, since the out of control situation is not a desirable one. This article proposes a methodology that helps to predict the main shifts (both in mean or scale), denoted as principal alarms, that can be expected in the process using the available in-control data. The outcome of the procedure can be used to improve the quality of the process in many ways. First, it helps to identify the weakest elements of the process, which will be responsible of most of the variability. This result can lead to quality engineers to implement some modifications in the process. This capacity to improve the process is especially appealing in the initial stages of the production process, where the engineers need to adjust the machinery.

Second, it allows to improve the SPC activity because: (i) it allows to choose the best type of control chart for the process, and (ii) it eases the search for responsible variables when an alarm already takes place.

The method is based on finding the independent factors that explain the variability of the process. Once these factors are estimated, a mean shift is introduced in them, one by one. The translation of these shifts, in terms of the original observed variables, is finally obtained. These shifts can be interpreted as a simulation of out of control situations provoked by each independent factor. The shifts provoked by the most important independent factors are called principal alarms.

These principal alarms can be represented graphically, allowing the practitioner to visualize the main expected deviation of the process. The principal alarms can also be used to select the best multivariate control chart. This comparison is made by estimating the ARL, of each competing chart, for shifts of different size. A resampling procedure for ARL estimation using principal alarms is proposed.

Acknowledgment

Authors are grateful to Alcalá Industrial SA for providing the case study, the data, and the permission for using Figure 1 (b). We are especially grateful to his Quality Engineer, Jose A. Delgado-Echague, for his comments and support. González's research is partly supported by the CICYT grant DPI2005-08018. Sánchez's research is partly supported by the grants CAM 06/HSE/0174/2004 and the CICYT grant SEJ2004-03303.

References

- [1] Alt, F.B. (1984). "Multivariate Quality Control," *Encyclopedia of Statistical Science*. Ed. S.Kotz and N. Johnson. New York: John Wiley, pp. 110-122.
- [2] Bajgier, S. (1992). "The Use of Bootstrapping to Construct Limits of Control Charts," *Proceedings of the Decision Science Institute*, San Diego, CA, pp. 1611-1613.
- [3] Beran, R. (1986). "Simulated Power Functions," *Annals of Statistics*, 14, pp. 151–173.
- [4] Crosier, R.B. (1988). "Multivariate Generalizations of Cumulative Sums Quality-Control Schemes," *Technometrics*, 30, pp. 291–303.
- [5] Doganaksoy, N., Faltin, F.W., and Tucker, W.T. (1991). "Identification of Out-of-Control Quality Characteristics in a Multivariate Manufacturing Environment," *Communications in Statistics-Theory and Methods*, 20, pp. 2775-2790.
- [6] Hawkins, D.M. (1991). "Multivariate Quality Control Based on Regression-Adjusted Variables," *Technometrics*, 33, pp. 61–75.
- [7] Hawkins, D.M. (1993). "Regression Adjustment for Variables in Multivariate Quality Control," *Journal of Quality Technology*, 25, pp. 170–182.
- [8] Hyvärinen, A., Karhunen, J., and Oja, E. (2001). *Independent Component Analysis*, John Wiley and Sons. New York.
- [9] Hotelling, H. (1947). "Multivariate Quality Control. Illustrated by the Air Testing of Sample Bombsights," in *Techniques of Statistical Analysis*. Eds. C. Eisenhart, M.W. Hastay, and W.A. Wallis. New York: McGraw-Hill, pp. 111-184.
- [10] Jones, L.A. and Woodall, W.H. (1998). "The Performance of Bootstrap Control Charts," *Journal of Quality Technology*, 30, pp. 362-375.
- [11] Liu, R. and Tang, J. (1996). "Control Charts for Dependent and Independent Measurements based on the Bootstrap," *Journal of the American Statistical Association*, 91, pp. 1694-1700.
- [12] Lowry, C.A., Woodall, W.H., Champ, C.W., and Rigdon, S.E. (1992). "A Multivariate Exponentially Weighted Moving Average Control Chart," *Technometrics*, 34, pp. 46–53.
- [13] Mason, R.L., Tracy, N.D., and Young, J.C. (1995). "The Composition of T^2 for Multivariate Control Chart Interpretation," *Journal of Quality Technology*, 27, pp. 99-108.
- [14] Mason, R.L., Tracy, N.D., and Young, J.C. (1997). "A Practical Approach for Interpreting Multivariate T^2 Control Chart Signals," *Journal of Quality Technology*, 29, pp. 396–406

- [15] Murphy, B.J. (1987). "Selecting Out-of-Control Variables with the T^2 Multivariate Quality Control Procedures," *The Statistician*, 36, pp. 571–583.
- [16] Seppala, T., Mokowitz, H., Plante, R., and Tang, J. (1995). "Statistical Process Control via the Subgroup Bootstrap," *Journal of Quality Technology*, 27, pp. 139-153.
- [17] Woodall, W.H., and Ncube, M.M. (1985). "Multivariate CUSUM Quality-Control Procedures," *Technometrics*, 27, pp. 285–292.